

The solution structure of a RNA pentadecamer comprising the anticodon loop and stem of yeast tRNA^{Phe}

A 500 MHz ¹H-n.m.r. study

G. Marius CLORE,* Angela M. GRONENBORN,* Edwin A. PIPER,† Larry W. McLAUGHLIN,‡
Erika GRAESER‡ and Jacques H. VAN BOOM§

*Division of Physical Biochemistry and †Division of Computing, National Institute for Medical Research, Mill Hill, London NW7 1AA, U.K., ‡Max-Planck Institut für Experimentelle Medizin, Abteilung Chemie, Hermann-Rein Strasse 3, D-3400 Göttingen, Federal Republic of Germany, and §Gorlaeus Laboratories, State University of Leiden, P.O. Box 9502, 2300 RA Leiden, The Netherlands

(Received 9 February 1984/Accepted 24 April 1984)

A 500 MHz ¹H-n.m.r. study on the semi-synthetic RNA pentadecamer 5'-r(C-A-G-A-C_m-U-G_m-A-A-Y-A-Ψ-m⁵C-U-G) comprising the anticodon loop and stem (residues 28–42) of yeast tRNA^{Phe} is presented. By using pre-steady-state nuclear-Overhauser-effect measurements all exchangeable and non-exchangeable base proton resonances, all H1' ribose resonances and all methyl proton resonances are assigned and over 70 intra- and inter-nucleotide interproton distances determined. From the distance data the solution structure of the pentadecamer is solved by model-building. It is shown that the pentadecamer adopts a hairpin-loop structure in solution with the loop in a 3'-stacked conformation. This structure is both qualitatively and quantitatively remarkably similar to that of the anticodon loop and stem found in the crystal structures of tRNA^{Phe} with an overall root-mean-square difference of 0.12 nm between the interproton distances determined by n.m.r. and X-ray crystallography. The hairpin-loop solution structure of the pentadecamer is very stable with a 'melting' temperature of 53°C in 500 mM-KCl, and the structural features responsible for this high stability are discussed. Interaction of the pentadecamer with the ribotrinucleoside diphosphate UpUpC, one of the codons for the amino acid phenylalanine, results only in minor perturbations in the structure of the pentadecamer, and the 3'-stacked conformation of the loop is preserved. The stability of the pentadecamer-UpUpC complex ($K \sim 2.5 \times 10^4 \text{ M}^{-1}$ at 0°C) is approximately an order of magnitude greater than that of the tRNA^{Phe}-UpUpC complex.

Most of our detailed knowledge of the structure of isolated tRNA species comes from single-crystal X-ray-diffraction studies (for reviews see Kim, 1981; Wright, 1982). In recent years, however, considerable progress has been made with the use of ¹H-n.m.r. spectroscopy to explore the solution structures of tRNA (Hare & Reid, 1982; Roy & Redfield, 1983; Heerschap *et al.*, 1983). These studies have principally made use of the nuclear Overhauser effect (n.O.e.) to demonstrate the

proximity of protons in space (Noggle & Schirmer, 1971). Because of the very large number of protons in tRNA (of the order of 1000) and the limited chemical-shift range of the proton, these studies have essentially been limited to the imino proton resonances, which occur in a relatively well-resolved and uncluttered region at the extreme low-field end of the ¹H-n.m.r. spectrum. Consequently, although such studies have provided extensive information on the overall structure of tRNAs in terms of secondary and tertiary hydrogen-bonding interactions involving imino protons, they cannot probe the details of the three-dimensional solution structures of individual loops and double-standed stems. This limitation can be overcome by making

Abbreviations used: n.O.e., nuclear Overhauser enhancement or effect; C_m, 2'-O-methylcytidine; G_m, 2'-O-methylguanosine; Y, wybutine; Ψ, pseudouridine; m⁵C, 5-methylcytidine.

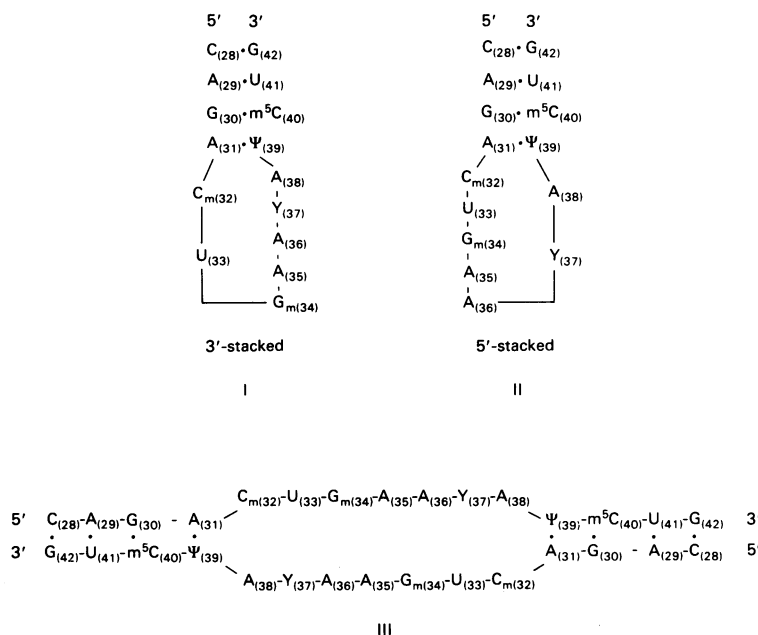


Fig. 1. Schematic representation of the three possible solution conformations of the RNA pentadecamer 5'-(C-A-G-A-C_m-U-G_m-A-A-Y-A-Ψ-m⁵C-U-G) comprising the anticodon loop and stem (residues 28–42) of tRNA^{Phe}. Structures I and II are hairpin-loop structures, whereas structure III is a bulge-duplex structure. The 3'-stacked conformation of loop shown in structure I is the configuration observed in the crystal structures of tRNA^{Phe}. The residues of the pentadecamer are numbered according to their numbering in tRNA^{Phe}.

use of synthetic oligoribonucleotides comprising specific portions of the tRNA molecule, thereby decreasing the spectral complexity and enabling n.O.e. measurements to be made on a larger number of proton types (namely imino, aromatic, ribose and methyl protons).

In the present investigation we have carried out a 500 MHz ^1H -n.m.r. study on the semi-synthetic RNA pentadecamer 5'-r(C-A-G-A-C_m-U-G_m-A-A-Y-A-Ψ-m⁵C-U-G) comprising the anticodon loop and stem (residues 28–42) of yeast tRNA^{Phe}. This pentadecamer can potentially adopt three structures in solution (see Fig. 1), of which the

distance data the solution structure of the pentadecamer is solved by model-building. It is shown that the solution structure exhibits great stability [T_m ('melting' temperature) $\sim 53^\circ\text{C}$] and is remarkably similar to that of the anticodon loop and stem found in the tRNA^{Phe} crystal structures. Moreover, it is only minimally perturbed on binding the ribotrinucleoside diphosphate UpUpC, the codon for the amino acid phenylalanine.

Experimental

The procedure used to synthesize the pentadecamer



hairpin-loop structure with the 3'-stacked loop conformation (structure I) is the one found in the crystal structures of tRNA^{Phe} (Jack *et al.*, 1976; Holbrook *et al.*, 1978). By means of extensive and systematic pre-steady-state proton-proton n.O.e. measurements, we have assigned all exchangeable and non-exchangeable base proton resonances, all H1' ribose resonances and all methyl and methylene resonances, and determined over 70 intra- and inter-nucleotide interproton distances. From the

was essentially as described by McLaughlin & Graeser (1982) and is therefore only be summarized here. Yeast tRNA^{Phe}, purified from bulk baker's-yeast tRNA by high-pressure liquid chromatography as described by Bischoff *et al.* (1983), was hydrolysed with ribonuclease T₁ (which is very specific for hydrolysis at guanosine) to yield a number of oligonucleotides, including the dodecamer



The latter was isolated by high-pressure liquid chromatography on a Zorbax-NH₂ column. A 5'-terminal phosphate was then added to the dodecamer by the use of ATP and polynucleotide kinase, and the phosphorylated dodecamer was isolated by anion-exchange chromatography on a Sephadex A-25 column and controlled for purity by high-pressure liquid chromatography on an ODS-Hypersil column. The ribotrinucleoside diphosphate rCpApG was synthesized from suitably protected nucleosides by using the 1-hydroxybenzotriazole phosphotriester approach (van der Marel *et al.*, 1981), de-blocked in a series of three steps, purified by anion-exchange chromatography on a Sephadex A-25 column and controlled for purity by high-pressure liquid chromatography on an APS-Hypersil column. The pentadecamer was then prepared from the trimer and dodecamer by the use of RNA ligase and isolated by preparative high-pressure liquid chromatography on an ODS-Hypersil column. After desalting and concentration of the purified pentadecamer, it was subjected to nucleoside and nucleoside 3'-phosphate analysis to verify its composition. The ribotrinucleoside diphosphate rUpUpC was synthesized and purified in the same way as rCpApG.

The samples for ¹H-n.m.r. were freeze-dried extensively from 99.6% ²H₂O and finally dissolved in either 99.96% ²H₂O or 90% H₂O/10% ²H₂O containing 500 mM-KCl, 10 mM-potassium phosphate buffer, pH*6.6 (meter reading uncorrected for the isotope effect on the glass electrode), and 0.02 mM-EDTA. Before use all glassware was heated at 200°C for 4 h to inactivate all possible traces of ribonuclease.

¹H-n.m.r. spectroscopy, n.O.e. measurements and model-building were carried out as described in the preceding paper (Gronenborn *et al.*, 1984). Chemical shifts are reported relative to 4,4-dimethylsilapentane-1-sulphonate.

Results and discussion

Resonance assignment

The 500 MHz ¹H-n.m.r. spectra of the RNA pentadecamer 5'-r(C-A-G-A-C_m-U-G_m-A-A-Y-A-Ψ-m⁵C-U-G) in 99.96% ²H₂O comprising only the non-exchangeable proton resonances and in 90% H₂O showing both the exchangeable and non-exchangeable proton resonances are illustrated in Figs. 2(a) and 3(a) respectively. The non-exchangeable proton resonances are labelled numerically and the exchangeable proton resonances alphabetically. Under the conditions of ionic strength (500 mM-KCl) and temperature (5°C and 0°C for the ²H₂O and H₂O spectra respectively) employed the four-base-pair stem of the pentadecamer is entirely double-stranded, as judged by the observa-

tion of four hydrogen-bonded imino proton resonances, peaks a-d, of the NH----N type (see Fig. 3a), and thermal-denaturation studies monitoring the temperature-dependence of the chemical shifts of the non-exchangeable proton resonances (*T*_m ~ 53°C; see subsection on thermal stability). All structural studies were carried out at these temperatures.

Resonance assignment was accomplished in a sequential manner by means of pre-steady-state n.O.e. measurements to demonstrate the proximity of protons in space (Wagner & Wüthrich, 1979). With the short irradiation times of 0.3 and 0.6 s used in this study, the pre-steady-state n.O.e. between two protons *i* and *j* of the pentadecamer is approximately proportional to *r*_{*ij*}⁻⁶, so that direct first-order n.O.e. values are only observable between protons separated by not more than 0.5 nm. The sequential assignment procedure was based on the principles and distance relationships described previously for right-handed double-helical DNA oligomers (Clare & Gronenborn, 1983, 1984; Reid *et al.*, 1983; Gronenborn *et al.*, 1984). Full use of this approach, however, could not be made for the RNA pentadecamer, as all the ribose proton resonances, with the exception of the H1' resonances, overlap in a narrow region of the spectrum between 4 and 5 p.p.m. and hence cannot be resolved individually even at a field strength of 500 MHz (see Fig. 2a). This contrasts with the situation in DNA oligomers, where the various deoxyribose resonance regions are well separated. Consequently, systematic measurements were restricted to irradiating in turn all methyl, H1' ribose and non-exchangeable base proton resonances (peaks 1-10 and 24-51; see Fig. 2a) and all exchangeable proton resonances (peaks a-k; see Fig. 3a). The assignment strategy was thus based on the observation of the following set of n.O.e. values between protons separated by not more than 0.5 nm (see Fig. 4 of Gronenborn *et al.*, 1984): (1) intranucleotide n.O.e. values involving the H1' ribose, base and methyl protons; (2) internucleotide n.O.e. values involving the base and H1' protons of adjacent nucleotides of the same strand that discriminate between the 5'→3' and 3'→5' directions in right-handed helices; (3) internucleotide n.O.e. values between adjacent H8/H6 base protons and adjacent H1' protons of the same strand; (4) intra-base-pair n.O.e. values between the imino and A(H2) protons of an A·U or A·Ψ base-pair, and between the imino and amino protons of a G·C base-pair; (5) inter-base-pair n.O.e. values involving imino, amino and A(H2) protons of adjacent base-pairs; (6) inter-residue n.O.e. values particular to the structure of the RNA pentadecamer, which involve the O2' methyl protons of residues C_{m(32)} and G_{m(34)}, the methyl

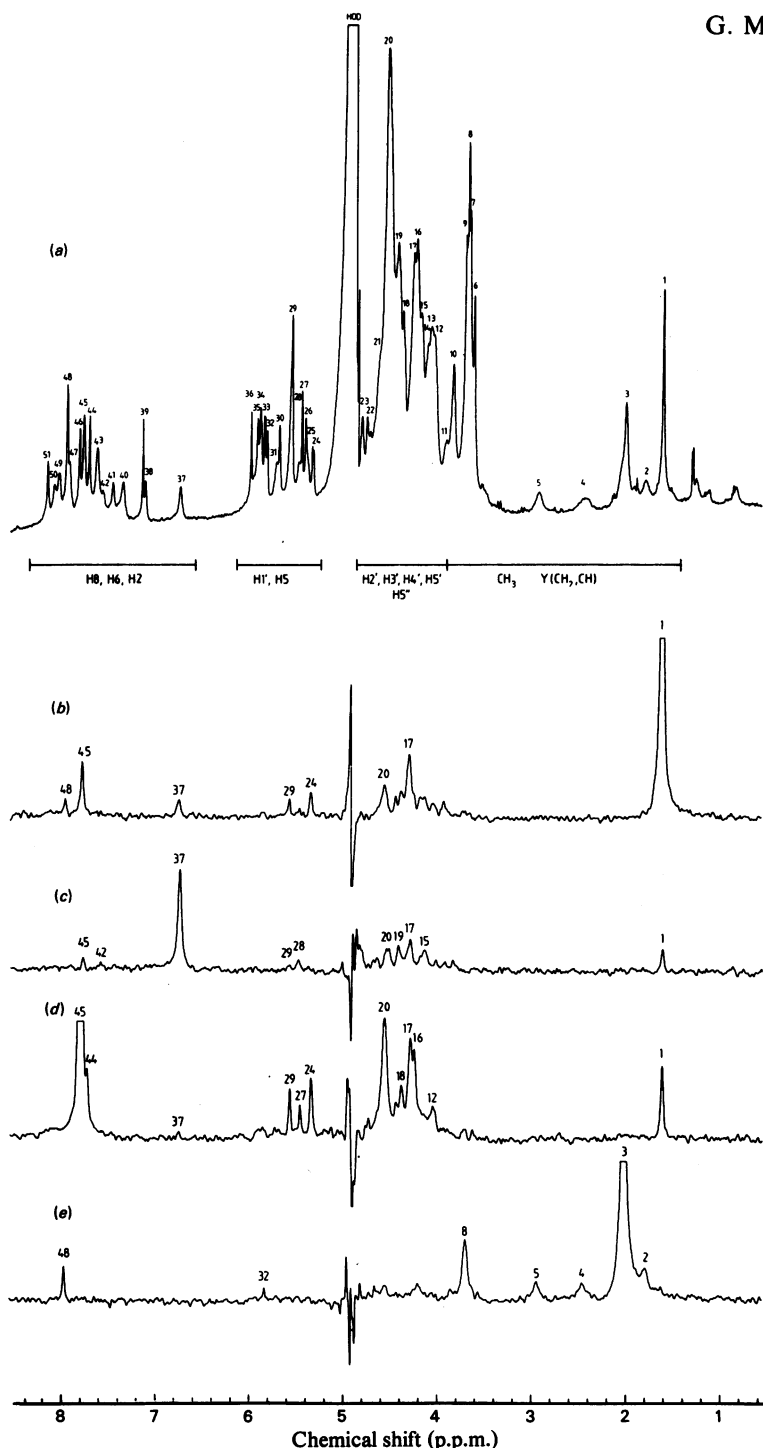


Fig. 2. Pre-steady-state n.O.e. measurements on the RNA pentadecamer in 99.96% $^2\text{H}_2\text{O}$ at 5°C (a) The 500 MHz ^1H -n.m.r. spectrum between 0 and 9 p.p.m. with all non-exchangeable proton resonances labelled 1–51 (unlabelled resonances are due to residual triethylammonium acetate and other minor impurities). (b)–(e) Difference spectra (off-resonance minus on-resonance pre-irradiation) after pre-saturation for 0.3 s of: (b) the $m^5\text{C}_{(40)}(\text{CH}_3)$ resonance (peak 1); (c) the $\Psi_{(39)}(\text{H6})$ resonance (peak 37); (d) the $m^5\text{C}_{(40)}(\text{H6})/\text{U}_{(41)}(\text{H6})$ resonance (peak 45); (e) the $\text{Y}_{(37)}(\text{C11CH}_3)$ resonance (peak 3). The assignments of the other peaks seen in the difference spectra are given in Table 1. The experimental conditions are: 0.9 mM-pentadecamer in 99.96% $^2\text{H}_2\text{O}$ containing 500 mM-KCl, 10 mM-potassium phosphate buffer, pH* 6.6 (meter reading uncorrected for the isotope effect on the glass electrode) and 0.02 mM-EDTA. Totals of 500 and 3200 transients were recorded for the reference and difference n.O.e. spectra respectively.

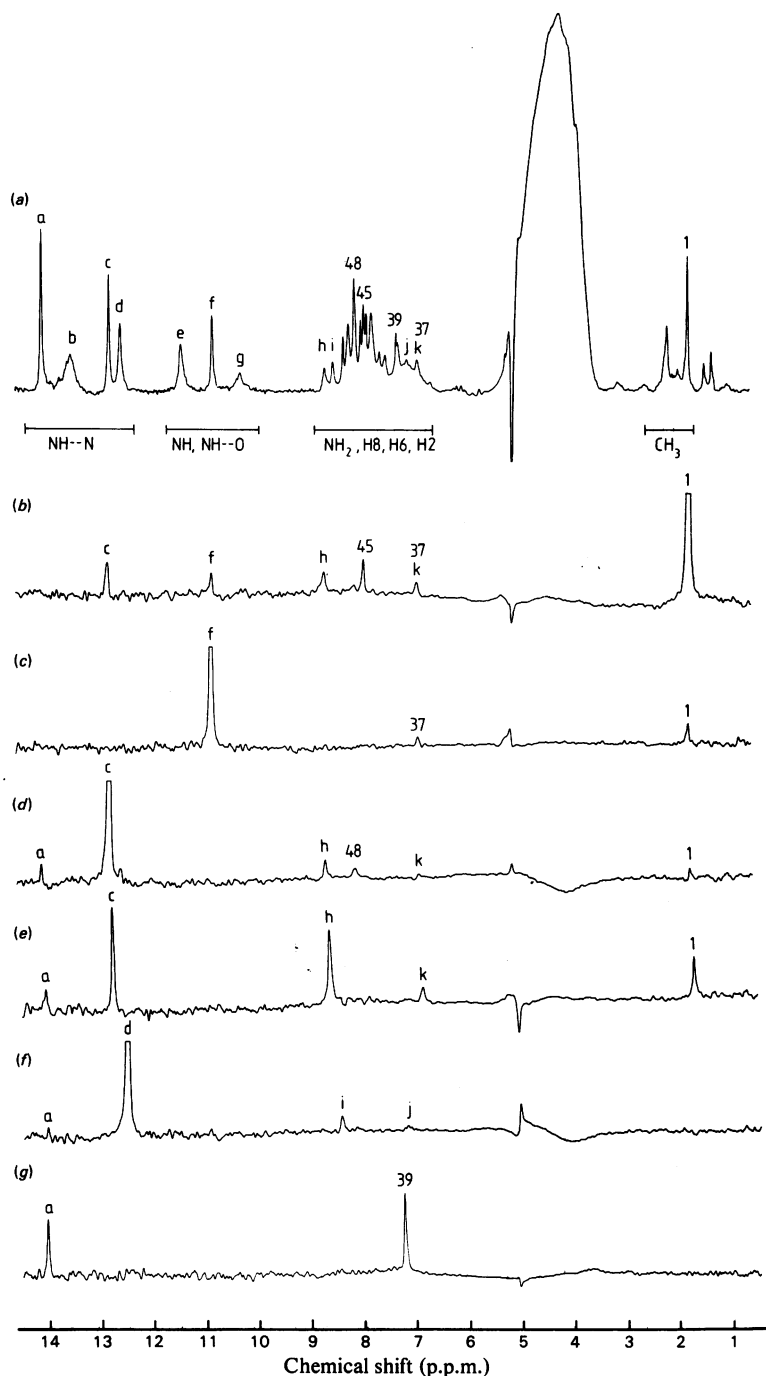


Fig. 3. *Pre-steady-state n.o.e. measurements on the RNA pentadecamer in 90% H₂O at 0°C*
 (a) The 500 MHz ¹H-n.m.r. spectra between 0 and 15 p.p.m. with the exchangeable imino proton resonances labelled a–g, the exchangeable amino proton resonance labelled h–k, and a few selected non-exchangeable proton resonances labelled as in Fig. 2. (b)–(g) Difference spectra (off-resonance minus on-resonance pre-irradiation) after pre-saturation of: (b) the m⁵C₍₄₀₎(CH₃) resonance (peak 1); (c) the Ψ₍₃₉₎(H1) resonance (peak f); (d) the G₍₃₀₎(H1) resonance (peak c); (e) the m⁵C₍₄₀₎(NH₂¹) resonance (peak h); (f) the G₍₄₂₎(H1) resonance (peak d); (g) the A₍₂₉₎(H2) resonance (peak 39). The assignments of the other peaks seen in the difference spectra are given in Table 1. The pre-saturation pulse was applied for 0.3 s in the case of the n.o.e. difference spectra (a)–(f) and for 0.6 s in the case of the n.o. difference spectrum (g). The experimental conditions are as in Fig. 2 except that the sample is in 90% H₂O/10% ²H₂O. Totals of 800 and 7200 transients were recorded for the reference and difference n.o.e. spectra respectively.

and methylene protons of the Y base, and a number of base and H1' protons within the loop.

From this large data set, the resonances of all exchangeable and non-exchangeable base protons, methyl protons and H1' ribose protons were assigned unambiguously without recourse to the crystal co-ordinates of the anticodon loop and stem of tRNA^{Phe}. The complete flow-chart of n.O.e.s observed for the RNA pentadecamer is shown in Fig. 4, the resonance assignments are given in Table 1, and the values of the n.O.e.s in Table 2. Some of the n.O.e.s outlined in the flow-chart are illustrated by the pre-steady-state n.O.e. difference spectra shown in Figs. 2 and 3.

One further point needs to be stated concerning the assignments of the exchangeable proton resonances e (11.29 p.p.m.) and g (10.70 p.p.m.), which are situated in the region of the spectrum characteristic of imino protons hydrogen-bonded to oxygen atoms and non-hydrogen-bonded imino protons. No n.O.e. values were observed involving either of these two resonances. By exclusion, however, they must be assigned to the loop imino protons of U₍₃₃₎ and G_{m(34)}. Resonance g is broad ($\Delta\nu_1 \sim 95$ Hz at 0°C) and only detectable up to 5°C, whereas resonance e is narrower ($\Delta\nu_1 \sim 35$ Hz at 0°C) and detectable up to 20°C. We therefore deduce that resonance e represents an imino proton hydrogen-bonded to an oxygen atom and

resonance g a non-hydrogen-bonded imino proton. On the basis of the internucleotide n.O.e. values involving the loop residues (see Figs. 4b and 4c), in particular those between the U₍₃₃₎(H1') and the A₍₃₅₎(H8) and A₍₃₆₎(H8) protons, and model-building, we conclude that peak e is the U₍₃₃₎(H3) resonance and peak g the G_{m(34)}(H1) resonance. This is discussed in detail below in the section dealing with the structure of the loop.

Solution structure of the RNA pentadecamer

Hairpin loop versus a bulge duplex. In principle the RNA pentadecamer could adopt either a hairpin-loop (structures I and II in Fig. 1) or bulge-duplex (structure III in Fig. 1) conformation. On thermodynamic grounds one would expect the hairpin-loop conformation to be considerably more stable than the bulge-duplex conformation, given that the number of residues in the loop is greater than four (Gralla & Crothers, 1973; Uhlenbeck *et al.*, 1973). These two conformations, however, can easily be distinguished on the basis of the correlation times calculated from the pre-steady-state intranucleotide n.O.e. values involving protons a fixed distance apart, in particular the H5/H1 and H6 protons of the pyrimidine base residues which are separated by 0.246 nm. In this manner, we obtain correlation times of 3.2 (± 0.2) and 1.8 (± 0.1) ns for these interproton vectors of

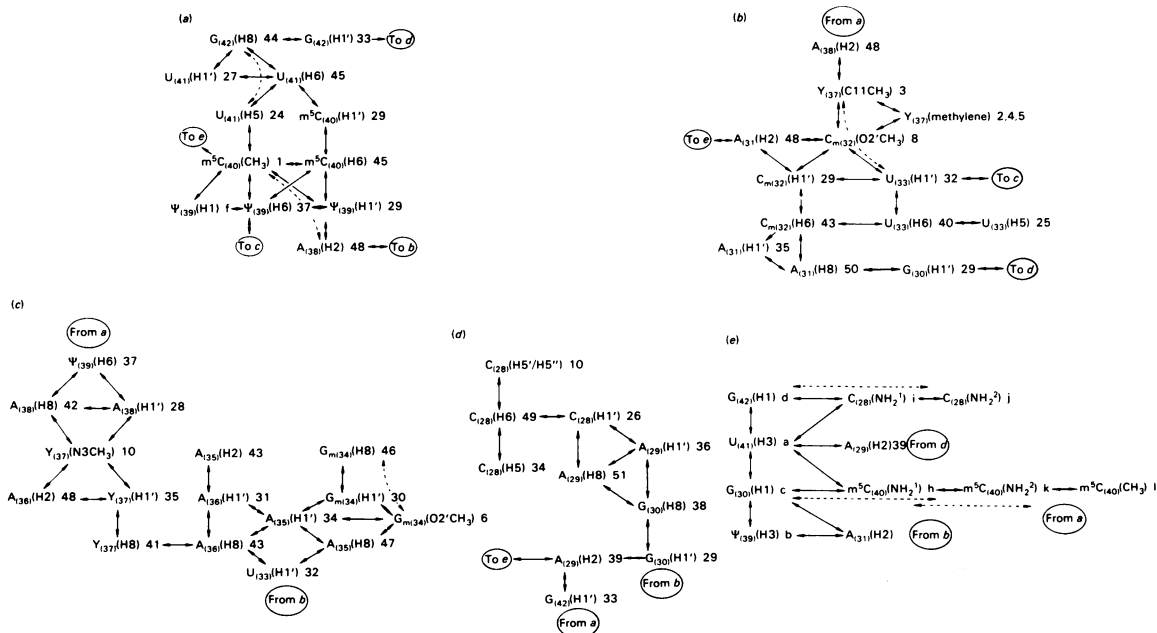


Fig. 4. Flow-chart of the observed n.O.e.s together with resonance assignments

The continuous lines (\longleftrightarrow) represent direct first-order n.O.e.s and the broken lines (\dashrightarrow) indicate indirect n.O.e.s

Table 1. Assignment of the proton resonances of the RNA pentadecamer

The chemical shifts are at 5°C. The other resonances assigned are as follows (peak no. in parentheses): $Y_{(37)}(C11CH_3)$, 2.01 p.p.m. (3); methylene protons of $Y_{(37)}$, 1.80 (2), 2.44 (4) and 2.92 p.p.m. (5); $Y_{(37)}(C16OOCH_3)$ and $Y_{(37)}(C21OOCH_3)$, 3.68 (7) and 3.72 p.p.m. (9); $Y_{(37)}(N3CH_3)$, 3.85 p.p.m. (10); $G_{m(34)}(O2'CH_3)$, 3.63 p.p.m. (6); $C_{m(32)}(O2'CH_3)$, 3.70 p.p.m. (8); $C_{(29)}(H5'/H5'')$, 3.85 p.p.m. (10). The $H2'$, $H3'$, $H4'$, $H5'$ and $H5''$ ribose resonances are located between 3.9 and 4.8 p.p.m. (peaks 11–23). The superscripts 1 and 2 in NH_2^1 and NH_2^2 refer to the hydrogen-bonded proton and the non-hydrogen-bonded proton respectively of the $-NH_2$ group. The amino protons are assigned to the C base rather than the G base, as the distance calculated from the values of the intra-base-pair n.O.e. between the imino and NH_2^1 amino protons (see Table 2b) is consistent with the idealized distance (Arnott & Hukins, 1972) between the G(H1) imino proton and the $C(NH_2^1)$ amino proton (approx. 0.26 nm) and is inconsistent with the distance between the G(H1) imino proton and the $C(NH_2^2)$ amino proton (approx. 0.23 nm).

Chemical shift [p.p.m. (peak no.)]

Residue	H8/H6	H5/C(5)CH ₃ /N(1)H	H2	H1'	NH	NH ₂ ¹	NH ₂ ²
$C_{(28)}$	8.05 (49)	5.91 (34)		5.43 (26)		8.39 (i)	7.12 (j)
$A_{(29)}$	7.17 (51)		7.16 (39)	6.00 (36)			
$G_{(30)}$	7.13 (38)			5.57 (29)	12.68 (c)		
$A_{(31)}$	8.10 (50)		7.97 (48)	5.94 (35)			
$C_{m(32)}$	7.65 (43)	5.57 (29)		5.57 (29)			
$U_{(33)}$	7.37 (40)	5.40 (25)		5.83 (32)	11.29 (e)		
$G_{m(34)}$	7.84 (46)			5.70 (30)	10.19 (g)		
$A_{(35)}$	7.94 (47)		7.65 (43)				
$A_{(36)}$	7.65 (43)		7.97 (48)	5.73 (31)			
$Y_{(37)}$	7.48 (41)			5.94 (35)			
$A_{(38)}$	7.58 (42)		7.97 (48)	5.50 (28)			
$\Psi_{(39)}$	6.75 (37)	10.70 (f)		5.57 (29)	13.43 (b)		
$m^5C_{(40)}$	7.79 (45)	1.62 (1)		5.57 (29)		8.56 (h)	6.78 (k)
$U_{(41)}$	7.79 (45)	5.34 (24)		5.47 (27)	13.96 (a)		
$G_{(42)}$	7.73 (44)			5.86 (33)	12.47 (d)		

the stem and loop pyrimidine base residues respectively. The difference between the correlation times of the loop and stem bases is easily explained as one would expect the amplitude of motion about the glycosidic bond to be larger for the loop residues than the base ones. These correlation times compare with values of 3 and 3.4 ns calculated from the data presented in the preceding paper (Gronenborn *et al.*, 1984) on a double-stranded DNA hexamer and octamer respectively under similar experimental conditions. As the correlation time is proportional to molecular mass, we conclude that the RNA pentadecamer adopts a hairpin-loop conformation under the experimental conditions of the present study.

Interproton distances. Interproton distances were calculated from the pre-steady-state n.O.e. data as described by Gronenborn *et al.* (1984). On the basis of the discussion in the preceding subsection, we have used the intranucleotide distance and n.O.e. between the H5(H1)–H6 pyrimidine protons as an internal reference, and assumed that the interproton vectors including only stem residues have the same correlation time as the reference vector of the stem pyrimidine bases, and that interproton vectors including any loop residue have the same correlation time as the reference vector of the loop

residue. With a relative error of not more than ± 0.15 in the measured n.O.e. values and an assumed error of ± 0.005 nm in the value of the reference distance, the error in the calculated interproton distances is not more than ± 0.02 nm. The complete set of 75 intra- and inter-nucleotide interproton distances determined from the n.O.e. measurements in this manner is given in Table 2.

An internal quality control on the n.m.r. distance determinations, with use of the H5–H6/H1–H6 intranucleotide interproton distances and n.O.e. values as internal references, is provided by the other two fixed internal distances present in the stem region of the pentadecamer: namely the $\langle r^{-6} \rangle^{-1/6}$ mean distance between the H6 and methyl protons of $m^5C_{(40)}$, which has an experimental and idealized value of approx. 0.25 nm, and the distance between the A(H2) and imino protons of the A·U and A· Ψ base-pairs, which has an experimental value of 0.26 nm compared with an idealized value of 0.28 nm (Arnott & Hukins, 1972).

Ribose and glycosidic-bond conformations. As only one sugar proton, namely the H1' proton, could be used in the n.O.e. measurements, assessment of the ribose conformation can only be made through the $J_{1',2'}$ coupling constant and the simple approximation $\% 3' \text{-endo} = 10 \times (10 - J_{1',2'})$ (Ezra *et al.*,

Table 2. Values of the pre-steady-state n.O.e. measurements (0.3 and 0.6 s pre-irradiation time) observed for the pentadecamer together with the interproton distances calculated from them and the corresponding interproton distances derived from the crystal structures of tRNA^{Phe}

(a) Intranucleotide n.O.e. values									
Base	N.O.e. (%)			Base	N.O.e. (%)			r_{ij} (nm)	
	0.3s	0.6s	N.m.r. ^a		0.3s	0.6s	N.m.r. ^a	X-ray ^b	
								Orthorhombic	Monoclinic
H1'-H8/H6				H1'-O(2')CH ₃					
C ₍₂₈₎	-6	-11	0.32	C _{m(32)}	-50	-70	0.20	0.19	0.21
A ₍₂₉₎	-10	-17	0.29	G _{m(34)}	-27	-40	0.22	0.21	0.24
G ₍₃₀₎	-3	-6	0.36	H1'-N(3)CH ₃	-9	-20	0.27	0.17	0.22
A ₍₃₁₎	-3	-4	0.38	Y ₍₃₇₎	-24	-45	0.25	0.24	0.25
C _{m(32)}	-2	-4	0.32 ^c	H6-H5/C(5)CH ₃ /N(1)H	-14	-27	0.25	0.24	0.25
U ₍₃₃₎	-8	-13	0.29	C ₍₂₈₎	-25	-40	0.25	0.24	0.25
G _{m(34)}	-2	-4	0.35	U ₍₃₃₎	-27	-50	0.25	0.26	0.28
A ₍₃₅₎	-4	-7	0.32	Y ₍₃₉₎	-25	-42	0.25	0.24	0.25
A ₍₃₆₎	g	-4	0.35	m ⁵ C ₍₄₀₎					
Y ₍₃₇₎ ^e	g	-3	0.40	U ₍₄₁₎					
A ₍₃₈₎	g	-8	0.34						
Y ₍₃₉₎	g	-8	0.34						
m ⁵ C ₍₄₀₎	-6	-8	0.34						
U ₍₄₁₎	-5	-8	0.34						
G ₍₄₂₎									

(b) N.O.e. values involving imino and amino protons participating in Watson-Crick hydrogen-bonding

Base	N.O.e. (%)			Base	N.O.e. (%)			r_{ij} (nm)	
	0.3s	0.6s	N.m.r. ^a		0.3s	0.6s	N.m.r. ^a	X-ray ^b	
								Orthorhombic	Monoclinic
Imino-imino				Amino-amino					
G ₍₄₂₎ (H1)-U ₍₄₁₎ (H3)	g	-2	0.42	C ₍₂₈₎ (NH ₂ ¹)-C ₍₂₈₎ (NH ₂ ²)	≥ -80	> -80	≤ 0.2	0.18	0.18
U ₍₄₁₎ (H3)-G ₍₃₀₎ (H1)	-2	-4	0.38	m ⁵ C ₍₄₀₎ (NH ₂ ¹)-m ⁵ C ₍₄₀₎ (NH ₂ ²)	≥ -80	> -80	≤ 0.2	0.18	0.18
G ₍₃₀₎ (H1)-Y ₍₃₉₎ (H3)	-2.5	-5	0.36	Amino-methyl	≥ -80	> -80	≤ 0.2	0.21	0.23
Imino-A(H2)				m ⁵ C ₍₄₀₎ (NH ₂ ²)-m ⁵ C ₍₄₀₎ (CH ₃)	≥ -80	> -80	≤ 0.2		
U ₍₄₁₎ (H3)-A ₍₂₉₎ (H2)	-18	-37	0.26						
Y ₍₃₉₎ (H3)-A ₍₃₁₎ (H2)	-18	-36	0.26						
G ₍₃₀₎ (H1)-A ₍₃₁₎ (H2)	-13	-28	0.28						
Imino-amino									
G ₍₄₂₎ (H1)-C ₍₂₈₎ (NH ₂ ¹)	-15	-30	0.27						
U ₍₄₁₎ (H3)-C ₍₂₈₎ (NH ₂ ¹)	-3	-5	0.36						
U ₍₄₁₎ (H3)-m ⁵ C ₍₄₀₎ (NH ₂ ¹)	-3	-5	0.36						
G ₍₃₀₎ (H1)-m ⁵ C ₍₄₀₎ (NH ₂ ¹)	-17	-36	0.26						

Table 2 (continued)
(c) Internucleotide involving non-exchangeable protons

Base	r_{ij} (nm)			N.O.e. (%)			r_{ij} (nm)		
	Base	N.m.r. ^a		0.3 s	0.6 s	N.O.e. (%)	N.m.r. ^a	X-ray ^b	
		Orthorhombic	Monoclinic					Orthorhombic	Monoclinic
C ₍₂₈₎	A ₍₃₅₎								
C ₍₂₈₎ (H1')-A ₍₂₉₎ (H8)	A ₍₃₅₎ (H1')-A ₍₃₆₎ (H1')	0.43	0.45	-10	-18	-2	0.35	0.50	0.46
C ₍₂₈₎ (H1')-A ₍₂₉₎ (H1')	A ₍₃₅₎ (H1')-A ₍₃₆₎ (H8)	0.50	0.52	-3	-5	-3	0.33	0.54	0.52
A ₍₂₉₎	A ₍₃₅₎ (H2)-A ₍₃₆₎ (H1')							0.28	0.50
A ₍₂₉₎ (H1')-G ₍₃₀₎ (H8)	A ₍₃₆₎	0.53	0.50	-8	-12	-2	0.35	0.45	0.45
A ₍₂₉₎ (H1')-G ₍₃₀₎ (H1')	A ₍₃₆₎ (H8)-Y ₍₃₇₎ (H8)	0.51	0.58	-3	-6	-6	0.29	0.42	0.46
A ₍₂₉₎ (H8)-G ₍₃₀₎ (H8)	A ₍₃₆₎ (H2)-Y ₍₃₇₎ (N3CH ₃)	0.51	0.52	-3	-7	-3	0.32	0.33	0.42
A ₍₂₉₎ (H2)-G ₍₄₂₎ (H1')	A ₍₃₆₎ (H2)-Y ₍₃₇₎ (H1')	0.40	0.35	-1.5	-3	-3			
G ₍₃₀₎	Y ₍₃₇₎								
G ₍₃₀₎ (H1')-A ₍₃₁₎ (H8)	Y ₍₃₇₎ (N3CH ₃)-A ₍₃₈₎ (H8)	0.49	0.51	-3	-5	-10	0.30	0.42	0.50
G ₍₃₀₎ (H1')-A ₍₃₁₎ (H1')	Y ₍₃₇₎ (N3CH ₃)-A ₍₃₈₎ (H1')	0.46	0.56	-2	-5	-7	0.28	0.26	0.38
A ₍₃₁₎	Y ₍₃₇₎ (C11CH ₃)-A ₍₃₈₎ (H2)						0.31	0.41	0.33
A ₍₃₁₎ (H1')-C _{m(32)} (H6)	A ₍₃₈₎	0.48	0.49	-2	-4	-6	0.29	0.44	0.47
A ₍₃₁₎ (H8)-C _{m(32)} (H6)	A ₍₃₈₎ (H1')-Y ₍₃₉₎ (H6)	0.47	0.41	-4	-10	-5	0.29	0.33	0.48
A ₍₃₁₎ (H2)-C _{m(32)} (H1')	A ₍₃₈₎ (H8)-Y ₍₃₉₎ (H6)	0.37	0.42	-13	-27	-13	0.25	0.35	0.34
A ₍₃₁₎ (H2)-C _{m(32)} (O2CH ₃)	A ₍₃₈₎ (H2)-Y ₍₃₉₎ (H1')	0.53	0.53	-18	-30	-13			
C _{m(32)}	Y ₍₃₉₎								
C _{m(32)} (H1')-U ₍₃₃₎ (H1')	Y ₍₃₉₎ (H1')-m ⁵ C ₍₄₀₎ (H6)	0.60	0.59	-3	-5	-8	0.30	0.46	0.47
C _{m(32)} (O2CH ₃)-U ₍₃₃₎ (H1')	Y ₍₃₉₎ (H1')-m ⁵ C ₍₄₀₎ (CH ₃)	0.60	0.43	-14	-20	-8	0.30	0.46	0.55
C _{m(32)} (O2CH ₃)-Y ₍₃₇₎ (C11CH ₃) ^b	Y ₍₃₉₎ (H6)-m ⁵ C ₍₄₀₎ (H6)	>0.7	>0.7	-16	-20	-8	0.30	0.54	0.49
C _{m(32)} (H6)-U ₍₃₃₎ (H5)	Y ₍₃₉₎ (H6)-m ⁵ C ₍₄₀₎ (CH ₃)	0.50	0.35	-5	-8	-16	0.27	0.39	0.36
C _{m(32)} (H6)-U ₍₃₃₎ (H6)	Y ₍₃₉₎ (H1)-m ⁵ C ₍₄₀₎ (CH ₃)	0.52	0.45	-5	-8	-6	0.33	0.41	0.35
C _{m(32)} (H5)-U ₍₃₃₎ (H6)	m ⁵ C ₍₄₀₎	0.54	0.55						
C _{m(32)} (H1')-U ₍₃₃₎ (H6)	m ⁵ C ₍₄₀₎ (H1')-U ₍₄₁₎ (H6)	0.60	0.45						
U ₍₃₃₎	m ⁵ C ₍₄₀₎ (CH ₃)-U ₍₄₁₎ (H5)								
U ₍₃₃₎ (H1')-A ₍₃₅₎ (H8)	U ₍₄₁₎	0.54	0.44	-5	-8	-9	0.30	0.44	0.49
U ₍₃₃₎ (H1')-A ₍₃₆₎ (H8)	U ₍₄₁₎ (H1')-G ₍₄₂₎ (H8)	0.38	0.55	-5	-10	-5	0.34	0.46	0.47
G _{m(34)}	U ₍₄₁₎ (H6)-G ₍₄₂₎ (H8)								
G _{m(34)} (H1')-A ₍₃₅₎ (H1')		0.63	0.60	-2	-5				
G _{m(34)} (H1')-A ₍₃₅₎ (H8)		0.45	0.51	-4	-8				
G _{m(34)} (O2CH ₃)-A ₍₃₅₎ (H1')		0.49	0.44	-7	-15				
G _{m(34)} (O2CH ₃)-A ₍₃₅₎ (H8)		0.43	0.49	-7	-11				

Table 2 (continued)

^a Interproton distances are calculated from the n.O.e. data as described by Gronenborn *et al.* (1984). In the distance calculations we have used the n.O.e. values obtained at 0.3 s irradiation except where these have values of less than -10% , in which case the value at 0.6 s irradiation is used. The fractional error $\Delta N/N$ in the n.O.e. measurements is not more than ± 0.15 , so that, assuming an error of ± 0.005 nm in the value of the reference distance, the error in the calculated interproton distances is not more than ± 0.02 nm. The reference distance used is the distance between the H5 (or H1) and H6 pyrimidine protons, which has a value of 0.246 nm on the basis of standard bond lengths and angles. Interproton distances involving only stem residues are calculated by using the mean H5-H6/H1-H6 n.O.e. value for the stem pyrimidine bases ($C_{(28)}$, $\Psi_{(39)}$ and $U_{(41)}$ at 0.3 s pre-irradiation time as the reference; interproton distances involving loop residues are calculated by using the H5-H6 n.O.e. value for $U_{(35)}$ at 0.3 s pre-irradiation time as the reference (see the text for discussion). All interproton distances given in the Table are rounded off to the second significant figure.

^b The interproton distances for the crystal structure of the anticodon loop and stem of tRNA^{Phe} are calculated from the refined co-ordinates of the orthorhombic (Sussman *et al.*, 1978) and monoclinic (Hingerty *et al.*, 1978) forms of yeast tRNA^{Phe} to which protons have been added with the use of standard bond lengths and angles. Distances involving methyl protons are $\langle r^{-6} \rangle^{-1/6}$ means calculated on the basis of free rotation of the methyl groups; this average is used instead of an arithmetic one in order to obtain a valid comparison with the interproton distances derived from n.O.e. measurements, which are all $\langle r^{-6} \rangle^{-1/6}$ means.

^c The resonances of $C_{m(32)}(H1')$ and $C_{m(32)}(H5)$ are superimposed in peak 29, so that the contribution of the n.O.e. between the $C_{m(32)}(H5)$ and $C_{m(32)}(H6)$ protons cannot be separated *a priori* from that between the $C_{m(32)}(H1')$ and $C_{m(32)}(H6)$ protons. The observed n.O.e. between peaks 43 [$C_{m(32)}(H6)$] and 29 [$C_{m(32)}(H1')$] and $C_{m(32)}(H5)$ have values of -20% and -40% at irradiation times of 0.3 s and 0.6 s respectively. Given that the distance between the H5 and H6 protons is fixed at 0.246 nm, we can assume that the n.O.e. values between the H5 and H6 protons of $C_{m(32)}$ will be the same as those between the H5 and H6 protons of $U_{(33)}$. On the basis of this assumption, n.O.e. values of -6% and -12% for irradiation times of 0.3 s and 0.6 s respectively between the H1' and H6 protons of $C_{m(32)}$ can be deduced, corresponding to a distance of 0.32 nm.

^d The resonances of $A_{(36)}(H8)$ and $A_{(35)}(H2)$ are superimposed in peak 43, so that the contribution of individual n.O.e. values to the observed n.O.e. between peak 43 and the $A_{(36)}(H1')$ resonance, peak 31 (-5% and -10% at irradiation times of 0.3 s and 0.6 s respectively), cannot be resolved.

^e In addition to the intranucleotide n.O.e. values for $Y_{(37)}$ given in the Table, large intranucleotide-n.O.e. values (over -40% at an irradiation time of 0.3 s) between the methylene protons of $Y_{(37)}$ (peaks 2, 4 and 5) and the $Y_{(37)}(C11CH_3)$ methyl protons are observed.

^f The resonances of $m^5C_{(40)}(H6)$ and $U_{(41)}(H6)$ are superimposed in peak 45 and those of $m^5C_{(40)}(H1')$ and $\Psi_{(39)}(H1')$ in peak 29, so that the contributions of individual n.O.e. values to the observed n.O.e. between peaks 29 and 45 (-20% and -40% at irradiation times of 0.3 s and 0.6 s respectively) cannot be resolved.

^g N.O.e. not detectable at 0.3 s irradiation time.

^h In addition to the n.O.e. between the $C_{m(32)}(O2'CH_3)$ and $Y_{(37)}(C11CH_3)$ protons, medium-sized n.O.e. values (approx. -10% at an irradiation time of 0.3 s) are observed between the $O2'CH_3$ methyl protons of $C_{m(32)}$ (peak 8) and the methylene protons of $Y_{(37)}$ (peaks 2, 4 and 5). In the crystal structure the distances between the $O2'CH_3$ methyl protons of $C_{m(32)}$ and the methylene protons of $Y_{(37)}$ are all not more than 0.5 nm.

ⁱ As the resonances of $C_{m(32)}(H1')$ and $C_{m(32)}(H5)$ are superimposed in peak 29, the contribution of the n.O.e. between these two protons and $U_{(33)}(H6)$, peak 40, cannot be resolved. The observed n.O.e. between peaks 29 and 40 has values of -9% and -12% respectively at irradiation times of 0.3 s and 0.6 s respectively.

^j The resonances of $U_{(41)}(H6)$ (peak 45) and $G_{(42)}(H8)$ (peak 44) are too close to enable one to quantify the n.O.e. between these two protons.

1977). Despite the relatively narrow linewidths ($\Delta\nu_1 \sim 2\text{Hz}$) for the H1' ribose resonances, no splittings could be detected for any of the resonances with the exception of the $G_{(42)}(\text{H1}')$ resonance (peak 33), which has a $J_{1',2'}$ coupling constant of approx. 4 Hz. We therefore conclude that all the sugar residues with the exception of that of $G_{(42)}$ are essentially pure 3'-endo, characteristic of A-RNA. The deviation from a pure 3'-endo conformation in the case of $G_{(42)}$ is attributable to an end effect. These findings are in complete agreement with the crystallographic data on tRNA^{Phe} (Jack *et al.*, 1976; Holbrook *et al.*, 1978).

The glycosidic-bond conformation can be deduced from the intranucleotide interproton distance between the H1' ribose and H8/H6 base protons by simple model-building. The syn and anti ranges for the glycosidic-bond torsion angle χ are $60 \pm 90^\circ$ and $240 \pm 90^\circ$ respectively, and $r_{\text{H1}'-\text{H8/H6}}$ has a maximum value of approx. 0.39 nm at $\chi = 240^\circ$ (anti) and a minimum value of approx. 0.25 nm at $\chi = 60^\circ$ (syn). In addition, each value of $r_{\text{H1}'-\text{H8/H6}}$ is compatible with two values of χ : $60^\circ < \chi_1 < 240^\circ$ and $\chi_2 = (240^\circ - \chi_1) + 240^\circ$. All the $r_{\text{H1}'-\text{H8/H6}}$ distances determined by n.O.e. measurements lie in the range 0.29–0.4 nm (see Table 2a), indicative of the anti conformation. Crystallographic studies on both tRNA^{Phe} and right-handed double-stranded oligonucleotides have shown that χ and the C4'–C3'-bond torsion angle δ are correlated, and for a value of δ in the 80–100° range, characteristic of a 3'-endo sugar conformation, χ lies in the low anti range (Holbrook *et al.*, 1978; Dickerson & Drew, 1981; Shakked *et al.*, 1983). On this basis we conclude that χ lies in the range 160–240°.

Structure of the stem region of the pentadecamer. The double-stranded nature of the stem is established by (i) the observation of four exchangeable imino proton resonances of the NH----N type arising from four Watson–Crick base-pairs, and (ii) the distance of approx. 0.26 nm determined by n.O.e. measurements between the pyrimidine bases (H3) and A(H2) protons of the $A_{(29)} \cdot U_{(41)}$ and $A_{(31)} \cdot \Psi_{(39)}$ base-pairs and between the G(H1) and C(NH₂¹) protons of the $G_{(42)} \cdot C_{(28)}$ and $G_{(30)} \cdot m^5C_{(40)}$ base-pairs (see Table 2b). The right-handed helical sense of the stem can be ascertained unambiguously from the finding that the $U_{(41)}(\text{H5})$ proton is closer to the $m^5C_{(40)}(\text{CH}_3)$ protons than is the $U_{(41)}(\text{H6})$ proton whereas the $\Psi_{(39)}(\text{H1})$ proton is further from the $m^5C_{(40)}(\text{CH}_3)$ protons than is the $\Psi_{(39)}(\text{H6})$ proton (see Table 2c). The average helical parameters of the stem region can then be determined by model-building with the use of the inter-residue interproton distances given in Tables 2(b) and 2(c) together with the glycosidic-bond (low anti) and ribose (3'-endo) conformations estab-

lished in the preceding subsection. This approach yields an average helical rise of approx. 0.25 nm, an average helical twist of approx. 33° and an average base-pair tilt of approx. 18°, in agreement with the single-crystal data on tRNA^{Phe} (Jack *et al.*, 1976; Hingerty *et al.*, 1978) and the fibre diffraction data on RNA polymers (Arnott *et al.*, 1973). Moreover, the n.m.r. interproton distance data are consistent with a separation of not more than 0.3 nm between the O2' atom of the 5'-residue and the O1' atom of the adjacent 3'-residue, thereby enabling the formation of O2'H----O1' hydrogen bonds, for all six possible cases within the stem.

Structure of the loop region of the pentadecamer. The first task in determining the structure of the loop region is to establish whether the loop is in the 3'- or 5'-stacked conformation represented by structures I and II respectively shown in Fig. 1. This is easily achieved on the basis of the following four observations: (i) direct n.O.e. connections between adjacent residues of the same strand are observed from $C_{(28)}$ through to $U_{(33)}$ (see Figs. 4b and 4d) and from $G_{m(34)}$ through to $G_{(42)}$ (see Figs. 4a and 4c); (ii) no n.O.e. connection could be observed between $U_{(33)}$ and $G_{m(34)}$; (iii) internucleotide n.O.e. values between the $U_{(33)}(\text{H1}')$ proton and the $A_{(35)}(\text{H8})$ and $A_{(36)}(\text{H8})$ protons are observed (see Fig. 4c); (iv) one of the loop imino proton resonances, resonance e (see Fig. 3a), probably represents an imino proton hydrogen-bonded to an oxygen atom on account of its narrow linewidth and detectability up to 20°C. These data are only compatible with the 3'-stacked conformation (Structure I, Fig. 1), in which there is a sharp turn at the junction of $U_{(33)}$ and $G_{m(34)}$ partly stabilized by a hydrogen bond between $U_{(33)}(\text{N3H})$ and $A_{(36)}(\text{O}_L\text{P})$ and the anticodon triplet bases $G_{m(34)}$, $A_{(35)}$ and $A_{(36)}$ are directed outwards in an approximately helical conformation. This is in complete agreement with the crystal structures of tRNA^{Phe} (Jack *et al.*, 1976; Holbrook *et al.*, 1978).

By using the same model-building approach as that used in the preceding subsection for the stem region, together with distance and stereochemical considerations to assign hydrogen bonds, we deduce the following additional structural features of the pentadecamer loop in solution (see Fig. 5).

(i) The sharp turn at the junction of $U_{(33)}$ and $G_{m(34)}$ is stabilized not only by a hydrogen bond between $U_{(33)}(\text{H3})$ and $A_{(36)}(\text{O}_L\text{P})$ but also by hydrogen bonds between $U_{(33)}(\text{O2}')$ and $A_{(35)}(\text{N6H})$ and between $U_{(33)}(\text{O2'H})$ and $A_{(35)}(\text{N7})$.

(ii) The core of the loop is stabilized by a hydrogen bond between $C_{m(32)}(\text{N4H})$ and $Y_{(37)}(\text{O6})$ and by hydrophobic interactions between the O2' methyl protons of $C_{m(32)}$ and the C11

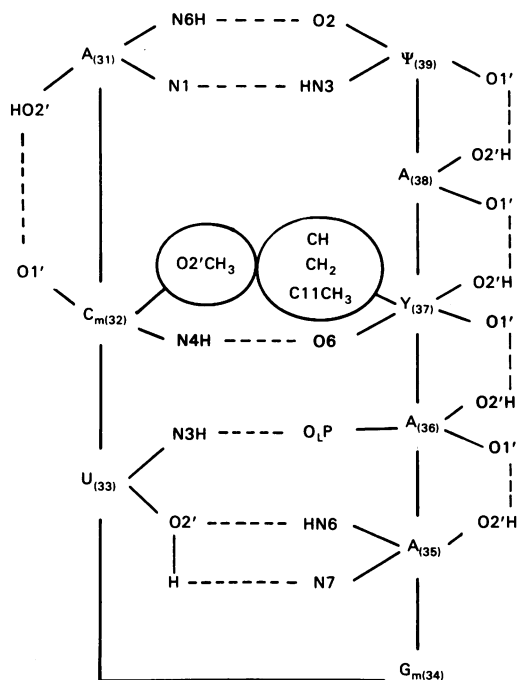


Fig. 5. Diagrammatic representation of the hydrogen-bonding and hydrophobic interactions stabilizing the 3'-stacked loop conformation of the pentadecamer in solution

See the text for discussion. Hydrogen bonds are represented by interrupted lines (----) and groups involved in hydrophobic interactions are encircled.

methyl and methylene protons of $Y_{(37)}$. The latter interactions involving the $C_{m(32)}(O2'CH_3)$ protons are not static, as rotation about the $C2'-O2'$ bond of $C_{m(32)}$ must occur to account for the n.o.e. values observed between the $C_{m(32)}(O2'CH_3)$ and $C_{m(32)}(H1')$, $Y_{(37)}(C11CH3)$ and $Y_{(37)}$ (methylene) protons (see Figs. 2e and 4b and Table 2c).

(iii) The two arms of the loop are additionally stabilized by O2'H----O1' hydrogen bonds between all the neighbouring residues with the exception of contacts involving U₍₃₃₎ and G_{m(34)}.

(iv) The only stabilizing interaction involving $G_{m(34)}$ is a stacking one with $A_{(35)}$. The $G_{m(34)}$ residue is therefore free to wobble as predicted by Crick's (1966) wobble hypothesis.

(v) There is rapid rotation about the C2'–O2' bond of $G_{m(34)}$. This is required to account for the n.O.e. values observed between the $G_{m(34)}(O2'CH_3)$ protons and the $G_{m(34)}(H1')$, $A_{(35)}(H1')$ and $A_{(35)}(H8)$ protons (see Fig. 4c and Table 2c).

Thermal stability of the hairpin-loop structure. Given the array of hydrogen bonds, outlined in the preceding subsection, stabilizing the loop region of

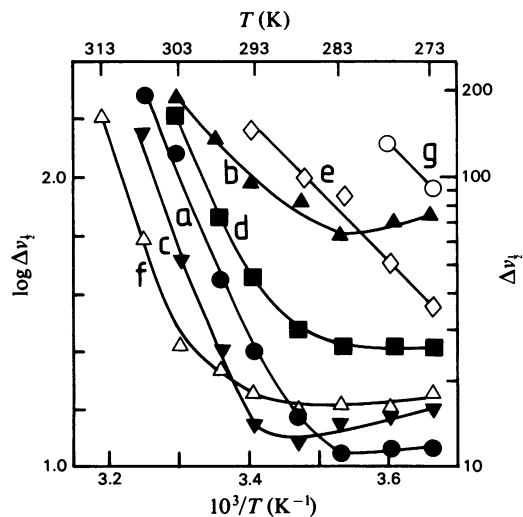


Fig. 6. Arrhenius plots of the temperature-dependence of the linewidths of the imino proton resonances
The labelling of the imino proton resonances is as in Fig. 3, and their assignments are given in Table 1. The experimental conditions are as in the legend to Fig. 3.

the pentadecamer, one would expect the pentadecamer to 'melt' in a co-operative manner at a relatively high temperature. This is indeed what is observed. The normalized 'melting' curves of five resonances that are easily monitored as a function of temperature, two from the stem [$G_{(30)}(H8)$ and $U_{(41)}(H5)$] and three from the loop [$C_{m(32)}(O2'CH_3)$, $U_{(33)}(H5)$ and $Y_{(37)}(CH_2)$], are superimposed, indicative of a co-operative 'melting' process for the stem and the loop regions. In 500 mM-KCl, the T_m , standard van't Hoff enthalpy of the 'melting' process and standard entropy have values of 53°C, 161 kJ·mol⁻¹ and 493 J·K⁻¹·mol⁻¹ respectively. These values should be compared with values of 21°C, 67 kJ·mol⁻¹ and 226 J·K⁻¹·mol⁻¹ in 1 M-NaCl for the oligoribonucleotide 5'-r(A₆C₆U₆), which has the most stable hairpin-loop structure of the series 5'-r(A₆C_nU₆) (Uhlenbeck *et al.*, 1973).

Imino proton exchange. Imino proton exchange rates provide a useful measure of hydrogen-bond stability. As the imino proton resonances do not shift over the temperature range over which they are detectable, their linewidth is given by $\Delta\nu_1 = (R_2 + k_{\text{ex}})/\pi$, where R_2 and k_{ex} are the spin-spin relaxation rate and overall imino proton exchange rate respectively. R_2 decreases slowly with increasing temperature as the pentadecamer tumbles more rapidly in solution, whereas k_{ex} increases as the temperature rises. As a result

Arrhenius plots of $\ln(\Delta\nu_i)$ versus $1/T$ are in general biphasic and the spin-spin-relaxation-rate contribution can be deduced from the low-temperature data.

On looking at the Arrhenius plots in Fig. 6, it is apparent that k_{ex} for the NH----N hydrogen-bonded imino protons increases in the order $G_{(30)}(\text{H1}) < U_{(41)}(\text{H3}) < G_{(42)}(\text{H1}) < \Psi_{(39)}(\text{H3})$, as expected from the sequence of the double-stranded stem region. It can also be seen that the $\Psi_{(39)}(\text{H1})$ exchanges even more slowly than the $G_{(30)}(\text{H1})$ imino proton. Given that $\Psi_{(39)}(\text{H1})$ is exposed in the major groove of the helical stem and as such is not shielded from solvent water, this observation indicates that $\Psi_{(35)}(\text{H1})$ is hydrogen-bonded to an oxygen atom. The only oxygen atom near enough to $\Psi_{(39)}(\text{H1})$ and in a favourable orientation to form a hydrogen bond is $\Psi_{(39)}(\text{O}_L\text{P})$. The effect of such hydrogen-bonding would be to diminish the motion about the glycosidic bond of $\Psi_{(39)}$. The loop imino protons exchange more rapidly than the other imino protons, and no spin-spin-relaxation-rate contribution to the linewidths of their resonances is apparent at 0°C . $G_{\text{m}(34)}(\text{H1})$ exchanges very rapidly ($k_{\text{ex}} \sim 300\text{s}^{-1}$ at 0°C), as expected from its exposed position, whereas $U_{(33)}(\text{H3})$ exchanges somewhat more slowly ($k_{\text{ex}} \sim 100\text{s}^{-1}$ at 0°C), on account of its participation in the U-turn loop hydrogen bond between $U_{(33)}(\text{H3})$ and $A_{(36)}(\text{O}_L\text{P})$.

The activation energies for exchange of the $\Psi_{(39)}(\text{H1})$ imino proton and the NH----N hydrogen-bonded imino protons, $G_{(30)}(\text{H1})$, $U_{(41)}(\text{H1})$ and $G_{(42)}(\text{H1})$, of the double-stranded stem are approximately equal with a value of approx. $150\text{kJ}\cdot\text{mol}^{-1}$. This value, which is comparable with the value of $161\text{kJ}\cdot\text{mol}^{-1}$ for the van't Hoff enthalpy of the 'melting' process, is significantly higher than that for the activation energy of individual base-pair opening (approx. $60\text{kJ}\cdot\text{mol}^{-1}$; Early *et al.*, 1981; Pardi *et al.*, 1982), and can therefore be interpreted as measuring the activation energy for the co-operative opening of

the helical stem and loop. The activation energies for exchange of the $\Psi_{(39)}(\text{H3})$ NH----N hydrogen-bonded imino proton and the two loop imino protons $G_{\text{m}(34)}(\text{H1})$ and $U_{(33)}(\text{H3})$, on the other hand, are quite low, with values around $50\text{kJ}\cdot\text{mol}^{-1}$.

Comparison of the solution and crystal structures. From the foregoing discussion on the solution structure of the loop and the stem regions of the pentadecamer, it is clear that the solution conformation of the pentadecamer is very similar to that of the anticodon loop and stem in the crystal structures of tRNA^{Phe}. In this respect, a comparison of the interproton distances determined from n.O.e. and crystallographic measurements is quite informative (see Table 2 and 3). In general, the agreement is good, particularly when it is borne in mind that the crystal structures were only solved to moderate resolution (0.25 and 0.27 nm for the monoclinic and orthorhombic forms respectively; Hingerty *et al.*, 1978; Sussman *et al.*, 1978), that the crystallographic thermal parameters for the anticodon loop and stem are large (with average mean square displacements of 0.055nm^2 in the orthorhombic form; Sussman *et al.*, 1978), and that small errors in co-ordinates can result in quite large errors (over 0.05 nm) in interproton distances. The overall root-mean-square difference between the n.m.r. and crystallographically determined interproton distances is 0.12 nm. Although these differences are not random, the n.m.r. distances being systematically shorter than the corresponding distances derived from the crystal structures, they can be accommodated by minor changes in backbone torsion angles and can be attributed to two main factors: (i) a somewhat tighter and more compact conformation in solution, possibly due to the absence of intermolecular interactions, particularly crystal packing forces, which have a major influence on the crystal structures of flexible regions such as the anticodon loop and stem of tRNA^{Phe}; (ii) the $\langle r^{-6} \rangle^{-1/6}$ mean nature of interproton distances determined by n.O.e. measure-

Table 3. Root-mean-square differences between the interproton distances determined by n.m.r. and X-ray crystallography for the anticodon loop and stem (residues 28–42) of tRNA^{Phe}

	Root-mean-square difference in interproton distances (nm)		
	N.m.r. – monoclinic	N.m.r. – orthorhombic	Monoclinic – orthorhombic
Intranucleotide ^a	0.04	0.05	0.03
Inter-residue (exchangeable protons) ^b	0.07	0.06	0.07
Inter-residue (non-exchangeable protons) ^c	0.16	0.16	0.08
Overall	0.12	0.12	0.07

^a Interproton distances given in Table 2(a).

^b Interproton distances given in Table 2(b).

^c Interproton distances given in Table 2(c).

ments. As the solution structure is dynamic and not static, this results in a set of n.m.r. distances weighted in favour of the fluctuations with the shortest interproton distances.

The most striking structural feature of both the solution and crystal structures is the 3'-stacked conformation of the loop, principally stabilized by the hydrogen bond between $U_{(33)}(H3)$ and $A_{(36)}(O_LP)$ (see Fig. 5). This hydrogen bond is quite strong, owing to the highly electronegative nature of the phosphate oxygen atom, and can be seen to play a crucial role in determining the structure of the loop (Kim & Sussman, 1976; Quigley & Rich, 1976; Jack *et al.*, 1976; Balasubramanian & Seetharamulu, 1983). In the case of the solution structure of the pentadecamer, the 3'-stacked loop conformation is additionally stabilized by four weaker interactions, namely hydrogen-bonding between $U_{(33)}(O2'H)$ and $A_{(35)}(N7)$, $A_{(35)}(N6H)$ and $U_{(33)}(O2')$, and $C_{m(32)}(N4H)$ and $Y_{(37)}(O6)$, and a hydrophobic interaction between the $O2'CH_3$ protons of $C_{m(32)}$ and the $C11CH_3$ and methylene protons of $Y_{(37)}$ (see Fig. 5). None of these weaker interactions is present in the orthorhombic crystal form of tRNA^{Phe} (Holbrook *et al.*, 1978), and only the first two are present in the monoclinic crystal form of tRNA^{Phe} (Jack *et al.*, 1976; Hingerty *et al.*, 1978). The absence of the hydrophobic interaction in the crystal structures is not surprising: in solution there is free rotation about the $C2'-O2'$ bond of $C_{m(32)}$ so that the $C_{m(32)}(O2'CH_3)$ protons can easily spend a portion of their time close (less than 0.3 nm) to the $Y_{(37)}(C11CH_3)$ and $Y_{(37)}(methylene)$ protons; in the crystal structures, however, the position of the $C_{m(32)}(O2'CH_3)$ protons is fixed approx. 0.4–0.5 nm from the $C_{m(32)}(H1')$ and $U_{(33)}(H1')$ protons and 0.5 nm or more from the $Y_{(37)}(C11CH_3)$ and $Y_{(37)}(methylene)$ protons (see Table 2c).

Functional properties of the RNA pentadecamer: interaction with 5'-r(UpUpC)

The equilibrium association constant for the interaction of the codon 5'-r(UpUpC) with tRNA^{Phe} is approximately three orders of magnitude larger than that expected for the interaction of the corresponding ribotrinucleoside diphosphates (Labuda & Pörschke, 1980; Pörschke & Labuda, 1982). This is thought to be due to the limited conformational flexibility imposed on the anticodon triplet by the overall conformation of the anticodon loop and the consequent almost perfect stacking of the bases of the codon and anticodon triplets in the tRNA^{Phe}-UpUpC complex (Clore *et al.*, 1984).

The only significant changes that can be detected in the ¹H-n.m.r. spectrum of the RNA pentadecamer on its interaction with UpUpC are:

(i) the disappearance of the non-hydrogen-bonded imino proton resonance of $G_{m(34)}$ (peak g); (ii) an upfield shift of approx. 35 Hz ($\equiv 0.07$ p.p.m.) in the position of the $\Psi_{(39)}(H3)$ hydrogen-bonded imino proton resonance (peak b); (iii) an increase (ranging from 7 to 15 Hz) in the linewidths of the other imino proton resonances, with the exception of the $G_{(42)}(H1)$ imino proton resonance (peak d), whose linewidth remains unaltered. No significant changes in either the positions or linewidths of the non-exchangeable proton resonances occur. Consequently, the increase in the linewidths of the imino proton resonances a, b, c, e and f (7, 16, 11, 8 and 15 Hz respectively) can be attributed to an increase in imino proton exchange rate ranging from 20 to 50 s⁻¹. In addition, no new hydrogen-bonded imino proton resonances are observed in the ¹H-n.m.r. spectrum of the RNA pentadecamer-UpUpC complex; this is not surprising, as the dissociation rate constant for the complex would be expected to be similar to that of the tRNA^{Phe}-UpUpC complex, namely approx. 2000 s⁻¹ (Labuda & Pörschke, 1980), and consequently the three hydrogen-bonded imino proton resonances of the codon-anticodon triplet would be broadened beyond detectability.

From the decrease in intensity of resonance g and the shift of peak b as a function of the amount of UpUpC added, the equilibrium association constant for the binding of UpUpC to the RNA pentadecamer is computed to be $2.6 (\pm 0.5) \times 10^4 M^{-1}$ at 0°C. This compares with a value of approx. $2 \times 10^3 M^{-1}$ at 7°C for the interaction of UpUpC with tRNA^{Phe} (Labuda & Pörschke, 1980).

From these data we conclude that the structure of the RNA pentadecamer remains essentially unaltered on its interaction with the triplet codon UpUpC and that the high stability of the RNA pentadecamer-UpUpC complex is attributable to a fairly rigid, restricted and highly stacked conformation of the anticodon triplet $G_{m(34)}-A_{(35)}-A_{(36)}$ imposed by the 3'-stacked structure of the pentadecamer loop stabilized by the hydrogen-bonding interactions discussed in the preceding subsection, particularly the $U_{(33)}(H3) \cdots A_{(36)}(O_LP)$ hydrogen bond (see Fig. 5).

Concluding remarks

In the present paper we have used the proton-proton n.O.e. to determine a large number of interproton distances from which the solution structure of the RNA pentadecamer 5'-r(C-A-G-A-C_m-U-G_m-A-A-Y-A-Ψ-m⁵C-U-G), comprising the anticodon loop and stem of tRNA^{Phe}, is elucidated. We have shown that in solution this RNA pentadecamer adopts a hairpin-loop structure with the

loop in the 3'-stacked conformation, which is preserved on binding of the codon UpUpC. Both the qualitative and quantitative aspects of this structure in solution have been examined and are found to be remarkably similar to those of the anticodon loop and stem in the crystal structures of tRNA^{Phe} (Jack *et al.*, 1976; Holbrook *et al.*, 1978). From the data presented in this paper it is clear that n.O.e. measurements on other synthetic RNA oligonucleotides comprising portions of the loops and stems of other biologically important higher-order RNA species, including other tRNA, rRNA, mRNA and plant RNA species, will provide a powerful method of structure determination where crystallographic data are not available. Moreover, if such data are combined with imino proton n.O.e. data characterizing the secondary and tertiary base-pairing in the intact species, then the complete elucidation of the three-dimensional solution structures of such larger higher-order RNA species may be attainable.

This work was supported by the Medical Research Council (G. M. C., A. M. G. and E. A. P.), the Lister Institute of Preventive Medicine (G. M. C.), the Max-Planck Gesellschaft (L. W. M.) and the Netherlands Organization for the Advancement of Pure Research ZWO (J. H. v. B.). G. M. C. is a Lister Institute Research Fellow. The ribotrinucleoside diphosphate rCpApG was prepared during a stay by L. W. M. in Leiden supported by an EMBO Short Term Fellowship. The n.m.r. spectra were recorded on the AM500 spectrometer of the Medical Research Council Biomedical NMR Centre at the National Institute for Medical Research.

References

- Arnott, S. & Hukins, D. W. L. (1972) *Biochem. Biophys. Res. Commun.* **47**, 1504–1509
- Arnott, S., Hukins, D. W. L., Dover, S. D., Fuller, W. & Hodgson, A. R. (1973) *J. Mol. Biol.* **81**, 107–122
- Balasubramanian, R. & Seetharamulu, P. (1983) *J. Theor. Biol.* **101**, 77–86
- Bischoff, R., Graeser, E. & McLaughlin, L. W. (1983) *J. Chromatogr.* **257**, 305–315
- Clore, G. M. & Gronenborn, A. M. (1983) *EMBO J.* **2**, 2109–2115
- Clore, G. M. & Gronenborn, A. M. (1984) *Eur. J. Biochem.* **141**, 119–129
- Clore, G. M., Gronenborn, A. M. & McLaughlin, L. W. (1984) *J. Mol. Biol.* **174**, 163–173
- Crick, F. H. C. (1966) *J. Mol. Biol.* **19**, 548–555
- Dickerson, R. E. & Drew, H. R. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7318–7322
- Early, T. A., Kearns, D. R., Hillen, W. & Wells, R. D. (1981) *Biochemistry* **20**, 3756–3764
- Ezra, F. S., Lee, C.-H., Kondo, N.S., Danyluk, S. S. & Sarma, R. H. (1977) *Biochemistry* **16**, 1977–1987
- Gralla, J. & Crothers, D. M. (1973) *J. Mol. Biol.* **73**, 497–511
- Gronenborn, A. M., Clore, G. M. & Kimber, B. J. (1984) *Biochem. J.* **221**, 723–736
- Hare, D. R. & Reid, B. R. (1982) *Biochemistry* **21**, 1835–1842
- Heerschap, A., Haasnoot, C. A. G. & Hilbers, C. W. (1983) *Nucleic Acids Res.* **11**, 4501–4520
- Hingerty, B., Brown, R. S. & Jack, A. (1978) *J. Mol. Biol.* **124**, 523–534
- Holbrook, S. R., Sussman, J. L., Warrant, R. W. & Kim, S. H. (1978) *J. Mol. Biol.* **123**, 631–660
- Jack, A., Ladner, J. E. & Klug, A. (1976) *J. Mol. Biol.* **108**, 619–649
- Kim, S. H. (1981) in *Topics in Nucleic Acid Structure*, part 1 (Neidle, S., ed.), pp. 81–112, Macmillan Press, London
- Kim, S. H. & Sussman, J. L. (1976) *Nature (London)* **260**, 645–646
- Labuda, D. & Pörschke, D. (1980) *Biochemistry* **19**, 3799–3805
- McLaughlin, L. W. & Graeser, E. (1982) *J. Liquid Chromatogr.* **5**, 2061–2077
- Noggle, J. H. & Schirmer, R. E. (1971) *The Nuclear Overhauser Effect: Chemical Applications*, Academic Press, New York
- Pardi, A., Morden, K. M., Patel, D. J. & Tinoco, I. (1982) *Biochemistry* **24**, 6567–6574
- Pörschke, D. & Labuda, D. (1982) *Biochemistry* **21**, 53–56
- Quigley, G. J. & Rich, A. (1976) *Science* **194**, 796–806
- Reid, D. G., Salisbury, S. A., Bellard, S. A., Shakked, Z. & Williams, D. H. (1983) *Biochemistry* **22**, 2019–2025
- Roy, S. & Redfield, A. G. (1983) *Biochemistry* **22**, 1386–1390
- Shakked, Z., Rabinovich, D., Kennard, O., Cruse, W. B. T., Salisbury, S. A. & Viswamitra, M. A. (1983) *J. Mol. Biol.* **166**, 183–201
- Sussman, J. L., Holbrook, S. R., Warrant, R. W., Church, G. M. & Kim, S. H. (1978) *J. Mol. Biol.* **123**, 607–630
- Uhlenbeck, O. C., Borer, P. N., Dengler, B. & Tinoco, I. (1973) *J. Mol. Biol.* **73**, 483–496
- van der Marel, G. A., van Boeckel, C. A. A., Wille, G. & van Boom, J. H. (1981) *Tetrahedron Lett.* no. 22, 3887–3890
- Wagner, G. & Wüthrich, K. (1979) *J. Magn. Reson.* **33**, 675–680
- Wright, H. T. (1982) in *Topics in Nucleic Acid Structure*, part 2 (Neidle, S., ed.), pp. 137–172, Macmillan Press, London